Journal of King Saud University – Computer and Information Sciences 35 (2023) 101568

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

# FedECG: A federated semi-supervised learning framework for electrocardiogram abnormalities prediction

Zuobin Ying [a], Guoyang Zhang [a,*], Zijie Pan [a], Chiawei Chu [a], Ximeng Liu [b]

[a] Faculty of Data Science, City University of Macau, Macau 999078, China
[b] College of Computer and Data Science, Fuzhou University, Fujian 350108, China

ARTICLE INFO

*Article history:*
Received 3 January 2023
Revised 31 March 2023
Accepted 21 April 2023
Available online 28 April 2023

*Keywords:*
Electrocardiogram
Cardiovascular diseases
Personal privacy
Label scarcity
Federated learning
Semi-supervised learning
Pseudo label
Data augmentation
ResNet
Non-independent and identically distributed

ABSTRACT

The soaring popularity of smart devices equipped with electrocardiograms (ECG) is driving a nationwide craze for predicting heart abnormalities. Smart ECG monitoring system has achieved significant success by training machine learning models on massive amounts of user data. However, three issues arise accordingly: 1) ECG data collected from various devices may display personal characteristic variations, leading to non-independent and identically distributed (non-i.i.d.) data. These differences can impact the accuracy and reliability of data analysis and interpretation; 2) Most ECG data on smart devices is unlabeled, and data labeling is resource-consuming as it requires heavy-loaded labeling from professionals; 3) While centralizing data for machine learning can address above issues like non-i.i.d. data and labeling difficulties, it may compromise personal privacy. To tackle these three issues, we introduce a novel federated semi-supervised learning (FSSL) framework named FedECG for ECG abnormalities prediction. Specifically, we adopt a pre-processing module to better utilize the ECG data. Next, we devise a novel model based on ResNet-9 in FSSL to accurately predict abnormal signals from ECG recordings. In addition, we incorporate pseudo-labeling and data augmentation techniques to enhance our implemented semi-supervised learning. We also develop a model aggregation algorithm to improve the model convergence performance in federated learning. Finally, we conduct simulations on a real-world dataset. Experiments demonstrate that FedECG obtains 94.8% accuracy with only 50% of the data labeled. FedECG achieved slightly lower accuracy than traditional centralized methods in ECG monitoring, with a 2% reduction. In contrast, FedECG outperforms the state-of-the-art distributed methods by about 3%. Moreover, FedECG can also support unlabeled data and preserve data privacy as well.



## 1. Introduction

In the past few years, the morbidity and mortality of cardiovascular diseases (CVDs) and other cardiac abnormalities have been increasing, cementing CVDs' position as the primary cause of death worldwide (Roth et al., 2020). The World Health Organization (WHO) estimates that 17.9 million deaths are associated with cardiac abnormalities annually (Kaptoge et al., 2019). Common cardiovascular risk factors, such as smoking, lack of exercise, poor diet, and excessive alcohol consumption, increase the likelihood of heart attacks, strokes, heart failure, arrhythmias, and other complications (Guddhur Jayadev and Bellary, 2022). Most patients are at risk of death at any time if their cardiac abnormalities remain undetected (Yusuf et al., 2020; Pandey et al., 2022). As a result, It is critical to predict cardiac abnormalities in time before heart attacks or strokes occur.

The rapid advancement of Internet-of-Things (IoTs) technology has created unprecedented possibilities in healthcare (Ying et al., 2022; Zhang et al., 2022). Smartwatches, fitness bracelets, and other health wearables now support medical-grade ECG functions (Philip et al., 2021). For example, the Apple Watch came with heart monitoring and was approved by the U.S. Food and Drug Administration (FDA) in 2017 (Wyatt et al., 2020). As smart devices are portable and wearable, they prove beneficial for users who wish to monitor ECG abnormalities over extended periods, particularly for the elderly and those too busy to visit the hospital. With

* Corresponding author.
*E-mail addresses:* zbying@cityu.mo (Z. Ying), D21091100356@cityu.mo (G. Zhang), zijiepan@protonmail.com (Z. Pan), cwchu@cityu.mo (C. Chu), snbnix@gmail.com (X. Liu).

Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2023.101568
1319-1578/

decreasing manufacturing costs, people are starting to buy ECG-equipped devices as a measure of health monitoring (Chatrati et al., 2022). .

ECG analysis often requires manual diagnosis, where doctors recognize and categorize various waveforms in the electrical signal to diagnose heart abnormalities. In the past decade, machine learning (ML) and artificial intelligence (AI) have been proposed in the medical field to assist doctors in diagnosing diseases by building classification models (Almotairi et al., 2023; Ying et al., 2021). Fig. 1 illustrates a common ECG monitoring system where the user's ECG-equipped device collects data without interruption (Baig et al., 2013; Yang et al., 2016). If the ECG recognition model in the device identifies an ECG abnormality, it notifies a private physician via SMS or calls an emergency number. However, this system poses a serious risk of data privacy leakage, leading to a reluctance to adopt such services. Meanwhile, several laws and regulations have been established to limit data availability and preserve user privacy. For instance, the EU General Data Protection Regulation (GDPR) has imposed strict regulations on data security and privacy protection (Regulation, 2016). Additionally, China and the US have increased their attention to data privacy (Zhang, 2018). In this background, it is almost impractical to integrate ECG data scattered on independent devices for model training.

ECG data collected by smart devices is typically unlabeled, which can be a challenge for professionals who may be overburdened with labeling and examining a vast amount of data (Kachuee et al., 2018). To address this issue, researchers often use the semi-supervised learning (SSL) paradigm (van and Hoos, 2020). This machine learning strategy involves mixing a small amount of labeled data with a large amount of unlabeled data during training. It enables the pre-trained model to be initialized with a small amount of labeled data and then updated continuously with unlabeled data. In Table 1, we summarize existing ECG monitoring works, including their architecture, learning strategy and model. However, most of these works focus on improving model accuracy while ignoring the issues of label scarcity and privacy preservation. Motivated by these works, we propose a novel federated semi-supervised learning framework for predicting ECG abnormalities (FedECG) to address these problems. Particularly, we present a realistic scenario whose framework is described in Fig. 2. In this scenario, the medical center server has a few ECG data labeled by professionals, while the user's smart devices have their unlabeled ECG data with personal privacy. In our FedECG implementation, we still face several challenges in predicting ECG abnormalities:

- **Data Pre-processing**: ECG data requires pre-processing to eliminate noise and ensure that it meets the input requirements for machine learning algorithms. Therefore, data pre-processing must be performed first before model training begins.
- **Privacy Protection**: Protecting user privacy is crucial, and medical servers must not have direct access to user data due to legal and regulatory restrictions. Thus lack of data for model training poses a challenge and personal privacy preserving must be considered while building models.
- **Label Scarcity**: ECG data collected by smart devices is unlabeled and requires expert annotation. However, labeling all the data is impractical due to the large volume, making label scarcity a critical challenge.
- **Non-i.i.d.**: Users may have different ECG data distributions and categories resulting in non-i.i.d. ECG situations. For example, some users may have normal beats, while others may have ventricular ectopic beats. This variation violates the assumption of i.i.d. typically used in machine learning.

To address the aforementioned challenges, we first demonstrate the ECG signal preprocessing process. In addition, we design a modified ResNet-9 model and propose a model weights clustering algorithm to improve the model's convergence. Specifically, this paper makes the following key contributions:

- **A federated semi-supervised learning framework for privacy-protecting ECG abnormalities prediction**: The framework considers the realistic situation where users' ECG data is noisy and unlabeled, and it employs a pre-processing process for ECG data, including signal denoising, segmentation, and transformation. Pseudo-labeling and data augmentation are utilized to label the users' data, and the federated learning approach is applied for global model training. This enables users to share model weights instead of their own data, which preserves privacy.
- **An improved ResNet-9 model for ECG abnormalities prediction**: We design a modified ResNet-9 model with eight convolutional layers and one fully connected layer. We overcome the gradient disappearance problem by incorporating residual blocks to achieve skipped connections. Considering the limited computing power of IoT devices, we set a smaller number of network layers and neurons. Experiments prove that the model has the advantage of small size and fast convergence.
- **A novel model weights clustering algorithm for improving model convergence performance**: We propose a model weights clustering algorithm to execute on the server. We adopt the K-means algorithm (Hartigan and Wong, 1979), which groups the client models based on the distribution of their weights and then aggregates them to the global model. Experimental results demonstrate that the algorithm can improve the model's accuracy by 2.7% and reduce about ten communication rounds in non-i.i.d. data distributions.

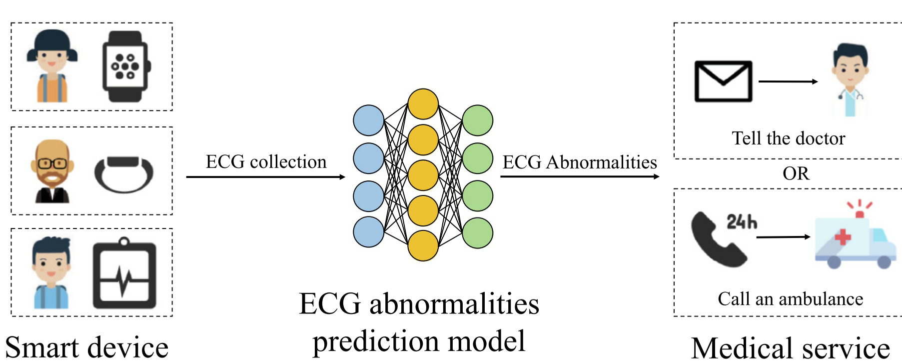


**Fig. 1.** ECG Monitoring System.

**Table 1**
Feature comparison of ECG monitoring works.

| Model | Architecture | Learning Strategy | Label Scarcity | Privacy Preserving |
|---|---|---|---|---|
| SVM (Kachuee et al., 2017) | Centralized | Supervised Learning | ✗ | ✗ |
| AdaBoost (Barstuğan and Ceylan, 2020) | Centralized | Supervised Learning | ✗ | ✗ |
| CNN (Wu et al., 2020) | Centralized | Supervised Learning | ✗ | ✗ |
| ResNet (Zhang et al., 2021) | Decentralized | Supervised Learning | ✗ | ✓ |
| CNN (Zhai et al., 2020) | Centralized | Semi-supervised Learning | ✓ | ✗ |
| Autoencoder (Rodrigues and Couto, 2021) | Centralized | Semi-supervised Learning | ✓ | ✗ |
| ResNet (Ours) | Decentralized | Semi-supervised Learning | ✓ | ✓ |

- **Extensive experiments and performance comparisons to evaluate the proposed FedECG framework's effectiveness**: The experiments consider various parameters, including model aggregation algorithm, number of clients, communication rounds, and levels of non-i.i.d. The results show that the proposed framework model achieves 94.8% accuracy at a 50% proportion of labeled data, considering a level of non-i.i.d. equal to 0.5. This demonstrates the framework's accurate ECG abnormality prediction on real-world datasets and its effectiveness on non-i.i.d. datasets.

## 2. Related work

### 2.1. ECG abnormalities monitoring methods

In recent years, smart wearable devices equipped with ECG monitoring have become increasingly popular in various fields, such as outdoor sports, health detection, etc. These devices continuously monitor users for any cardiac abnormalities by collecting ECG data (Ali et al., 2021; Zhang et al., 2022). To predict and detect cardiac abnormalities, existing models often employ machine learning and artificial intelligence techniques. For example, Afkhami et al. (2016) adopted a Gaussian mixture model to fit ECG data and obtain feature vectors for training a decision tree model. Kachuee et al. (2017) used support vector machines (SVM) to solve the problem of continuous and noninvasive blood pressure estimation in ECG monitoring systems. It is widely acknowledged that feature engineering is often required for most machine learning problems (El-Sappagh et al., 2022). Recent findings (Singh et al., 2022; Hussein et al., 2022) show that deep neural networks can extract features directly from the medical data. For instance, Wu et al. (2020) proposed a convolutional neural network (CNN) with 12 resilient and efficient layers for distinguishing

**Table 2**
Categories of heartbeats defined by AAMI.

| Class | Heartbeat |
|---|---|
| Normal | Normal beat<br>Left bundle branch block beat<br>Right bundle branch block beat |
| Ventricular ectopic | Premature ventricular beat<br>Ventricular escape beat |
| Supraventricular ectopic | Atrial premature beat<br>Aberrated atrial premature beat<br>Nodal premature beat<br>Supraventricular premature beat |
| Fusion | Fusion of Ventricular and normal beat |
| Unknown | Paced beat<br>Fusion of paced and normal beat<br>Unclassifiable beat |

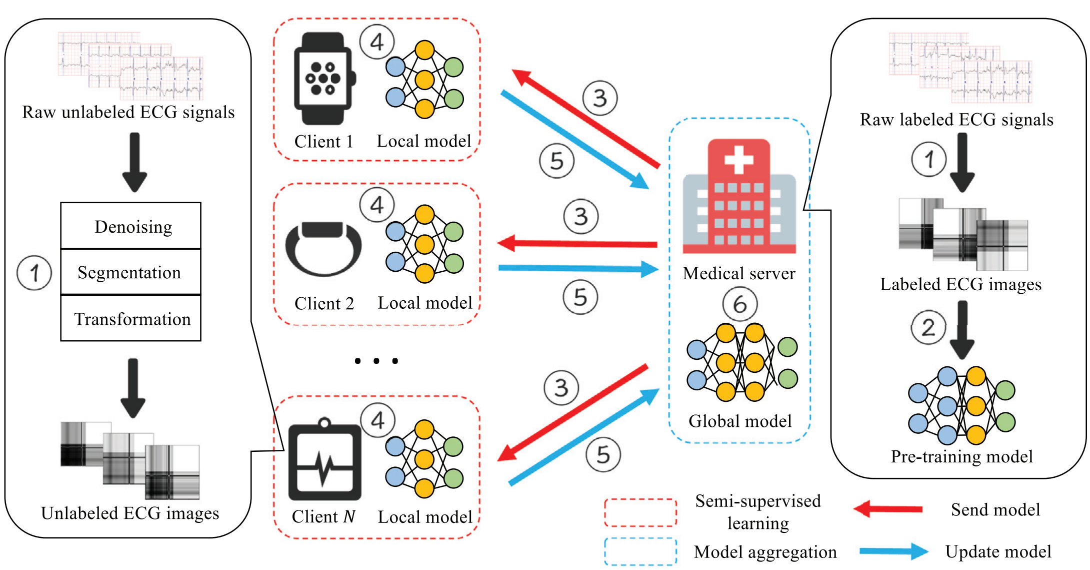


**Fig. 2.** Overview of proposed FedECG framework. The FedECG framework process works in six steps: 1) The central medical server and clients separately pre-process their own raw datasets; 2) The medical server uses labeled ECG data to pre-train the global model; 3) The medical server sends the trained global model parameters to the clients, such as smart watches, bracelets, at-home ECG monitors, etc.; 4) The clients perform semi-supervised learning technologies with local unlabeled data using the global model to generate a new local model; 5) The clients upload their local model parameters to the medical server; 6) The server aggregates multiple local model parameters to get a new global model.

heartbeat types in an ECG dataset. Ali et al. (2020) presented a new heartbeat classification boosting algorithm based on ensemble deep learning and extracted valuable features using the information gain method. Undoubtedly, machine algorithms have helped to solve the problem of limited medical resources and improved the accuracy of ECG abnormalities prediction.

The development of ECG AI models relies on public ECG datasets like PhysioNet (Moody et al., 2001). The Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Arrhythmia database (Moody and Mark, 2001) is recognized as the most widely distributed dataset for ECG signal processing and arrhythmia identification among the free datasets. As shown in Table 2, the five heartbeat categories classified by MIT-BIH are in line with those specified by the Association for Advancement of Medical Instrumentation (AAMI) (Stergiou et al., 2018). In this work, we use the MIT-BIH dataset to simulate our actual data in the FedECG framework. Since all ECG signals are annotated by two cardiologists, the data labels are more accurate.

### 2.2. Labeled data scarcity in ECG monitoring

With the increasing popularity of ECG-equipped devices, more and more ECG data is being collected from users. However, most ECG data collected from smart device sensors are unlabeled, and machine learning techniques require a large quantity of labeled data to train models (Mehari and Strodthoff, 2022). Unfortunately, the annotation task is expensive and time-consuming.

To address the issue of data scarcity, especially with unbalanced datasets, semi-supervised learning is becoming a popular solution (Guo et al., 2022). The most popular semi-supervised methods for ECG include self-learning, data augmentation, active learning, and so on. For example, Sarkar and Etemad (2020) proposed an ECG emotion recognition system using pseudo-labeling technology. Wagner et al. (2018) addressed the lack of labeled data in the ECG data streams through temporal label propagation. Furthermore, data augmentation in ECG has been tackled using generative adversarial network (GAN) models to produce simulated data that is more authentic than that acquired by other methodologies (Golany et al., 2021). However, GAN models require a large amount of data, which goes against the need for more data for ECG monitoring. Drawing inspiration from previous works, we consider techniques such as pseudo-labeling and data augmentation to address the labeled data scarcity in ECG.

### 2.3. Privacy-preserving in ECG monitoring

The ECG monitoring system is designed to create models that can predict ECG abnormalities. However, this requires the aggregation of ECG data on a cloud server, which can compromise user privacy. To address this issue, researchers have developed privacy-preserving methods using cryptographic algorithms. For example, Huang et al. (2019) presented a practical solution to validate patients with noisy ECG signals while maintaining privacy. Şahinbaş and Catak (2021) applied the secure multi-party computing (MPC) technology to preserve ECG data analysis. Kocabas and Soyata (2020) used the homomorphic encryption (HE) for privacy-preserving healthcare cloud computing. However, these methods reduce the accuracy of model learning and increase the time consumption of learning due to the complexity of algorithms.

Federated learning has emerged as a promising solution for privacy-preserving ECG monitoring in IoT devices (Wagan et al., 2022). According to a recent review by Yang et al. (2019), federated learning methods can be categorized into three types: vertical, horizontal, and transfer. Our proposed FedECG framework adopts a horizontal federated learning approach, where participating ECG-equipped devices share the same feature space but have distinct samples. Zhang et al. (2020) proposed a machine learning algorithm based on federated learning that does not require all local ECG data to be collected on an external platform for centralized learning. Research by Raza et al. (2022) has shown that the accuracy of federated learning solutions for ECG monitoring is comparable to traditional methods. In conclusion, federated learning is an optimal privacy-preserving approach for ECG monitoring systems.

## 3. Preliminaries

The section provides a quick overview of federated learning and semi-supervised learning algorithms.

### 3.1. Federated averaging algorithm

Federated learning enables users to construct machine learning models while keeping their data local. The Federated Averaging (FedAvg) algorithm introduced by McMahan et al. (2017) serves as a foundational element in federated learning environments. This algorithm combines local stochastic gradient descent for each client with a server that performs iterative model averaging. Let $N$ denotes the set of clients and $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$ indicate the dataset for each client. The FedAvg algorithm proceeds as follows:

**Algorithm 1** FedAvg Algorithm

**ServerModelAggregation:**
1: Initialize global model weight $\omega_0$
2: **for** each communication round $t = 1, 2, 3, \ldots$ **do**
3: $N_t \leftarrow$ (Sample a set of $N$ clients)
4: **for** each client $n \in N_t$ **in parallel do**
5: $\omega^n_{t+1} \leftarrow$ ClientModelUpdating$(n, \omega_t)$
6: **end for**
7: $\omega_{t+1} \leftarrow \frac{1}{|N_t|}\sum_{n\in N_t}\omega^n_{t+1}$
8: **end for**
**ClientModelUpdating$(n, \omega)$:**
9: $\eta$ is the learning rate, $\ell$ is the loss function.
10: **for** each epoch $i = 1, 2, \ldots, I$ **do**
11: $\omega \leftarrow \omega - \eta\nabla\ell(\omega)$
12: **end for**
13: send $\omega$ to server

- **Initializing Parameters**: In each communication round $t$, the server selects a group of clients $n$ randomly from $N$ client sets to engage in the federated learning process. Then the server transmits the initial global model parameters $w_t$ to the $n$ clients.
- **Local Model Training**: When each of $n$ clients receives the global model parameters from the server, they train local models with their data $X_n$ for $i$ epochs. The goals of clients are to reduce the local loss function $\ell_i$:

$$\arg\min_{\omega^n\in\mathbb{R}} L_n(\omega^n) = \frac{1}{|X_n|}\sum_{(x_i,y_i)\in X_n}\ell_i(y_i, f_k(x_i;\omega^n)), \tag{1}$$

where $X_n$ represents the local dataset, the local dataset includes independent variables $x_i$ and dependent variables $y_i$, and $w^n$ is the local model parameter. Clients, after training models, upload their local model parameters to the server.
- **Global Model Aggregation**: The server adopts the FedAvg to gather the received model parameters and get a novel global model $w_{t+1}$ for the following rounds:

$$\omega_{t+1} = \frac{1}{|N_t|}\sum_{n\in N_t}\omega^n_t, \tag{2}$$

repeating the mentioned procedures before the global loss function reaches its minimum value means that the global model $w_{t+1}$ is converged.

The complete procedure of the FedAvg algorithm is displayed in the Algorithm 1. Federated learning tasks often incorporate differential privacy or homomorphic encryption during the parameter exchange and gradient aggregation steps. These techniques serve as robust tools for safeguarding the learning process and have been extensively researched and utilized in federated learning.

### 3.2. Pseudo-labeling

Pseudo-labeling is a fundamental concept in semi-supervised learning (SSL) proposed by Lee et al. (2013) to enhance the accuracy of machine learning models when labeled data is limited. Pseudo-labeling employs a trained model to annotate the unlabelled data and then retrains the model on both the original labeled data and newly pseudo-labeled data.

- **Pseudo Label Generation**: In the process of predicting unlabeled data, pseudo-labeling selects the classification with the highest probability from the predicted outcomes as the label for the unlabeled data. The following definition is provided:

$$y_i' = \begin{cases} 1, & \arg\max f(x_i) \\ 0, & \text{others} \end{cases} \tag{3}$$

where $f(x_i)$ represents the output of the neural network, $y_i'$ is the pseudo label of $x_i$.
- **Minimizing the Loss Function**: One method for converting unsupervised learning into a standard objective function term is by applying entropy regularization. This process treats unlabeled data with pseudo labels as if they were labeled data and employs a loss function, such as the cross-entropy loss function, to assess the error magnitude. In practice, the total loss can be expressed as follows:

$$\text{Totalloss} = \text{Labelloss} + \alpha * \text{Unlabelloss}, \tag{4}$$

where $\alpha$ is used to control the contribution of unlabeld data to the overall loss.

### 3.3. Data augmentation

Data augmentation is a popular machine learning approach to address the problem of overfitting and limited labeled data (Shorten and Khoshgoftaar, 2019). This technique produces new data by fine-tuning the existing data or adding noise, ensuring that the labels of the new data and the original data are the same. This approach can create additional data to compensate for the shortage of labeled data, which can improve model performance.

There are two primary types of data augmentation methods: weak augmentation (Zheng et al., 2021) and strong augmentation (Yuan et al., 2021). Weak augmentation involves processing the data without changing the content of the data itself. Common weak augmentation methods include panning, rotating, flipping, etc. (Bin et al., 2020). Although the data may not change significantly for humans, these methods can modify the arrangement or position of the data, which can be helpful for machine learning models. Strong augmentation adds noise or perturbation to the data to change the content of the data itself. Common methods of strong augmentation are RandAugmentation (Choi et al., 2022) and CTAugmentation (Sandfort et al., 2019). However, it can be useful in some instances, such as when the model must learn to recognize objects in a cluttered or noisy environment.

In this work, we perform a weak and a strong augmentation separately with unlabelled data, and the new data generated can assist in the pseudo-labeling process.

## 4. Methodology

In this section, we outline the specific technical implementations of our federated semi-supervised learning (FSSL) framework for ECG analysis. Firstly, we propose approaches for signal denoising, segmentation, and transformation. Secondly, we describe the structure of our ECG abnormalities predicting model. Thirdly, we provide details on the federated learning averaging and aggregation algorithms employed in our framework. Finally, we develop a semi-supervised learning method to address issues related to label scarcity.

### 4.1. ECG data processing

The ECG data processing involves several steps, including signal denoising, signal segmentation, and signal-to-image transformation. As shown in Fig. 3, the ECG signals are initially denoised and segmented into individual ECG beats. Each beat is subsequently transformed into an image by plotting its time–amplitude waveform and saving it as an image format. This conversion enables further analysis using image-based techniques.

#### 4.1.1. ECG signal denoising

During the process of collecting ECG data, various types of noise frequently distort the signals. These noises include myoelectric noise, muscle artifacts, and instrument noise, which usually fall within the frequency range of 1 Hz. To enhance the accuracy of ECG analysis, it is essential to eliminate these noises from the signals. One commonly used approach is the wavelet transform, which has been shown to be effective in studies (Tuncer et al., 2019; Yang and Wang, 2022). The wavelet transform is an effective method because it can analyze signal information in both the time and frequency domains. It is represented as follows:

$$WT(s, \imath) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(x) * \Psi\left(\frac{x - \imath}{s}\right) dx, \tag{5}$$

where $s$ represents the scale variable and $\imath$ represents the translation variable. The $s$ and $\imath$ denote the scaling and translation of the wavelet function. The scale parameter is inversely proportional to frequency, while the translation parameter $\imath$ represents time. The wavelet transform is effective in removing high-frequency noise from the ECG signals while preserving the essential features of the ECG waveform. This is beneficial for the training of the ECG abnormalities prediction model.

#### 4.1.2. ECG signal segmentation

ECG analysis involves the detection of the amplitude, shape, and relative position of the waveforms. Since collected ECG signals typically contain multiple ECG beats, it is necessary to segment the ECG into individual beats to enable targeted training.

The continuous ECG signals can be split into multiple beats based on the time interval between the previous R wave peak and the next R wave peak in the ECG waveform, which represents a heartbeat cycle. All extracted beats have the same length, which is essential for input into subsequent processing stages. If an ECG beat is not long enough, it can be padded with zeros to a fixed size. A single ECG beat is defined as follows:

$$T(Rpeak(x-1)) \leqslant T(x) \leqslant T(Rpeak(x+1)), \tag{6}$$

which $T(x)$ indicates a segment of ECG. Because the wavelet transform method mentioned above makes the ECG signal more periodic,

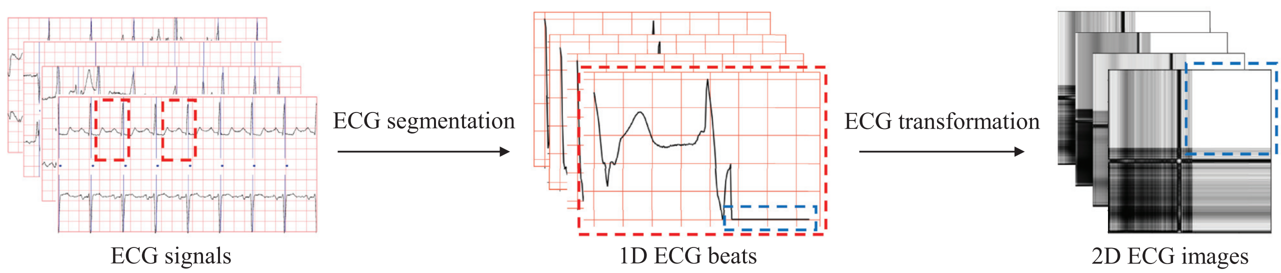


**Fig. 3.** The graphical representation of ECG data processing.

it is possible to segment the different signals with a fixed time interval.

*4.1.3. Signal-to-image transformation*

Although deep learning methods such as 1D-CNN (Sabor et al., 2022), long short-term memory network (LSTM) (Saadatnejad et al., 2019), gated recurrent unit (GRU) (Ebrahimi et al., 2020), and others can directly handle one-dimensional (1D) data, most existing approaches primarily address two-dimensional (2D) structured data. Deep learning methods can automatically discover the most important features from raw image data, which has been proven effective in various computer vision tasks. However, in the case of ECG data, previous works have experimentally demonstrated that a 2D representation of ECG provides more accurate heartbeat classification than a 1D representation (Li et al., 2022).

To exploit the benefits of visually interpreting ECG signals, we encode ECG signals as images using a method called Gramian Angular Fields (GAF) (Wang and Oates, 2015). This method allows the 1D signal to be represented as 2D data while preserving the time information on the time series. The basic procedures of the GAF formula are shown below:

- **Numerical Scaling**: First, considering that $\mathscr{M}$ is the set of $n$ segmented ECG signals, it can be written as $\mathscr{M} = \{m_1, m_2, \ldots, m_i, m_j, \ldots, m_n\}$. We normalize $\mathscr{M}$ into [0,1] with a Min–Max scaler using the following formula:

$$\bar{m}_i = \frac{m_i - \min(\mathscr{M})}{\max(\mathscr{M}) - \min(\mathscr{M})}, m_i \in \mathscr{M}. \tag{7}$$

- **Coordinate Conversion**: Next, we encode the value as angular cosine and the timestamp as a radius to map the rescaled ECG signal $\bar{\mathscr{M}}$. The following formulas are used to explain this coding method:

$$\alpha = arccos(\bar{m}_i), \tag{8}$$

$$R_i = \frac{t_i}{C}, \tag{9}$$

where $t_i$ is the time stamp, $R$ represents the spatial domain affiliation, and $C$ is a constant used to modify how widely the angular coordinate system spreads. In short, we obtain the delimiting points by dividing the interval [0, 1] into $C$ equivalent portions and associating these points with the time series.
- **Trigonometric Transformation**: Finally, we adopt the angular perspective by factoring the total or variance of the trigonometric functions of samples to demonstrate correlations among several timestamps. The following equations explain the Gramian Angular Field (GAF) method used in this paper:

$$\begin{aligned} \text{GASF} &= cos(\alpha_i + \alpha_j) \\ &= \bar{A} \cdot \bar{A}' - \sqrt{B - \bar{A}^2} \cdot \sqrt{B - \bar{A}'^2}, \end{aligned} \tag{10}$$

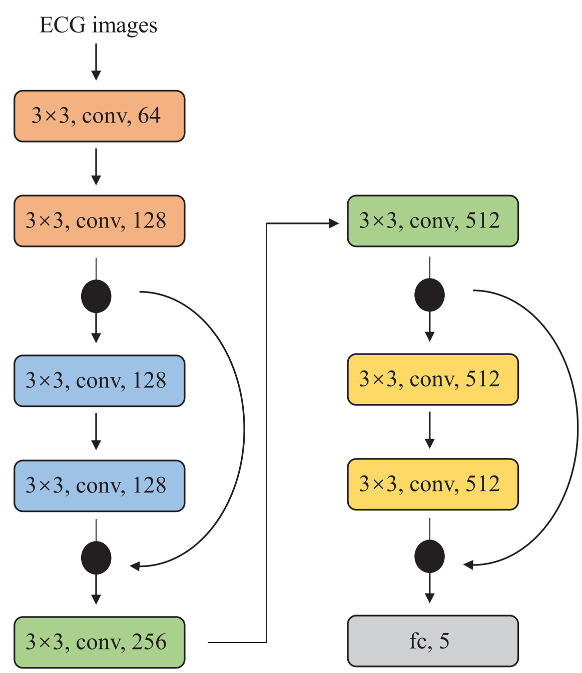


**Fig. 4.** The architecture of the proposed network.

$$\begin{aligned} \text{GADF} &= sin(\alpha_i - \alpha_j) \\ &= \sqrt{B - \bar{A}'^2} \cdot \bar{A} - \bar{A}' \cdot \sqrt{B - \bar{A}^2}, \end{aligned} \tag{11}$$

where GASF stands for Gramian Angular Summation Fields, and GADF stands for Gramian Angular Summation Fields, $A$ and $A'$ denote different row vectors, and $B$ is the unit row vector. We note that the GAF extends the operation of the traditional inner product in a penalized version. Its matrix density is sparse, which can better explain the temporal correlation.

*4.2. ECG abnormalities predicting model*

In this work, we introduce a novel ResNet-9 model for predicting ECG abnormalities. The residual neural network (ResNet) was first introduced by He et al. (2016) in 2015 and has since become widely used in various applications such as detection, segmentation, and recognition due to its simplicity and effectiveness. As depicted in Fig. 4, our proposed ResNet-9 model is lightweight and consists of three key components: a convolutional layer, a residual block, and a fully connected layer.

*4.2.1. Convolutional layer*

The main responsibility of the convolutional layer is to extract the main features of the input data. To accomplish this, we use eight convolutional layers in our network and add a pooling layer after them. Additionally, we apply batch normalization (BN) before the pooling layer to improve the neural network training, increase the convergence rate, and preserve network robustness.

We employ a 1D convolution of length 64 to extract features from the initial data. The length of the subsequent convolutional layer is increased continuously to obtain high-level spatial features. We set the kernel size of each convolutional layer to $(3 \times 3)$ and their strides to be 1.

#### 4.2.2. Residual block

To address the problem of gradient disappearance or explosion, ResNet introduces the concept of residual blocks, which use skip connections to connect a layer to other layers by skipping some intermediate layers. The residual block formula is expressed as follows:

$$b_i = F(a_i, w_i) + h(a_i), \tag{12}$$

$$a_{i+1} = f(b_i), \tag{13}$$

where $a_i$ and $a_{i+1}$ indicate the input and output vector of the residual block $i$, respectively. $w_i$ denotes the convolution operation, $F$ is the residual function representing the learned residuals, $h(a_i)$ is used to adjust the dimension, and $f$ is the activation function, typically the ReLU function.

By preserving the information from the previous input during the new intake process, the network avoids the need for additional parameters and computations. This significantly enhances the model's training efficiency. Moreover, this structure effectively addresses the degradation issue that can arise as the model's depth increases.

#### 4.2.3. Fully connected layer

The fully connected layer acts as a classifier in the entire network. It maps the extracted features of the output into a 1D vector. We use one fully connected layer for the classification of the five types of heartbeats. For multi-class classification, we use the softmax activation function in the final layer. The softmax function maps the network output to a probability distribution over the different classes. The formula for the softmax function is as follows:

$$\text{Softmax }(g_i) = \frac{e^{g_i}}{\sum_{k=1}^{n} e^{g_k}}, \text{for } i = 1, \ldots, n, \tag{14}$$

where $n$ denotes the number of classes, $g_i$ is the logit for class $i$, and $Softmax(g_i)$ is the predicted probability for class $i$. The softmax function ensures that the predicted probabilities sum to one, and each predicted probability is between 0 and 1, which makes it suitable for multi-class classification.

### 4.3. Semi-supervised learning

The clients only hold unlabeled data on their local devices, while the server holds labeled data. In this work, we adopt pseudo-labeling and data augmentation methods to train a local model with unlabeled data for the clients. The main steps of the semi-supervised learning procedure are illustrated in Fig. 5 and described as follows:

- **Step 1**: The server performs regular supervised learning with labeled data. It has access to a labeled data batch $L = \{(p_b, q_b), b \in (1, \ldots, B)\}$, where $p_b$ represents the input data and $q_b$ is the corresponding ground-truth label. The model function $f$ is trained using the supervised loss function $\ell_s$, which is expressed as the average cross-entropy loss over all labeled data points:

$$\ell_s = \frac{1}{B} \sum_{b=1}^{B} \mathrm{H}(q_b, f(p_b)), \tag{15}$$

where H is the cross-entropy loss function, and $f(p_b)$ represents the predicted output of the model.
- **Step 2**: Clients without labeled data use weak augmentation to obtain pseudo labels. They apply the standard rotation and shift operations to the ECG images and pass them through the model for training. If the model's prediction score is above a certain threshold, the corresponding classification is used as a pseudo label.
- **Step 3**: Once the client obtains the pseudo labels, it performs strong augmentation with the unlabeled data. The strong augmentation aims to calculate the cross-entropy loss from the pseudo labels and the features. The unlabeled data batch is denoted as $U = \{u_b, b \in (1, \ldots, u_B)\}$, and the corresponding pseudo labels are denoted as $\hat{q}_b$. The following formulas are used for this step:

$$\hat{q}_b = \mathrm{argmax}(f(\hat{u_b}) \geqslant \tau), \tag{16}$$

$$\ell_u = \frac{1}{u_B} \sum_{b=1}^{u_B} \mathrm{H}(\hat{q}_b, f(u_b)), \tag{17}$$

where $\ell_u$ is the unsupervised loss function, $\hat{u_b}$ is the unlabeled data after weak augmentation, $\tau$ is the threshold value for retaining pseudo labels.

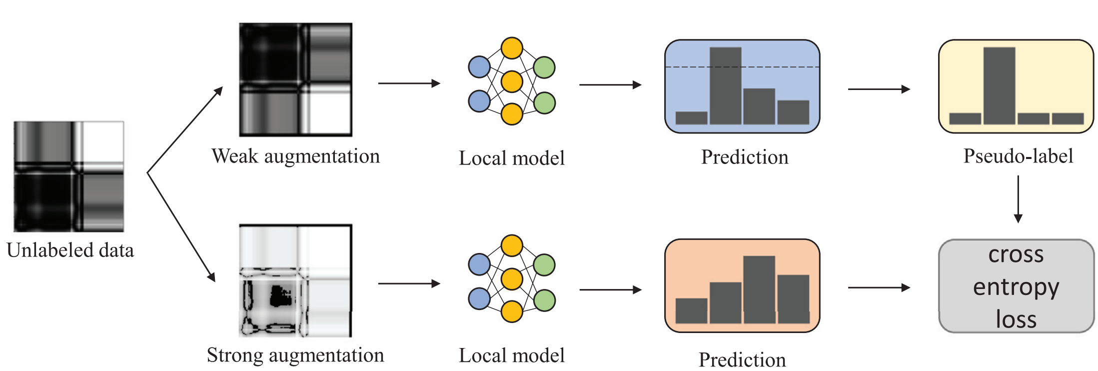


**Fig. 5.** Diagram of Semi-Supervised Learning.

- **Step 4**: Finally, the server combines its loss function with the clients' cross-entropy loss functions to update the whole global model by backward gradient propagation of the total loss function:

$$\text{Loss} = \ell_s + \lambda \ell_u, \tag{18}$$

where $\lambda$ is the hyperparameter weight for balancing the contribution of supervised loss and unsupervised loss.

### 4.4. Federated learning

The foundational component of the FedECG framework is federated learning, which enables multiple participating clients to train a global model without compromising their privacy. We make an improvement on the clustering step in the FedAvg algorithm as a compensation strategy to address the issue of non-independent and identically distributed (non-i.i.d.) data.

**Algorithm 2** Model Weights Clustering Algorithm for Non-i.i.d. Data

Initialize global model weight $\omega_0$
**for** each communication round $t$ = 1, 2, 3, ... **do**
  $N_t \leftarrow$ (Sample a set of $N$ clients)
  **for** each client $n \in N_t$ **in parallel do**
    $\omega^n_{t+1} \leftarrow$ ClientModelUpdating($n$, $\omega_t$)
  **end for**
  $C_{t+1} \leftarrow$ ClusteringAlgorithm($\omega_{t+1}$)
  **for** each cluster $c \in C_{t+1}$ **in parallel do**
    $m \leftarrow$ (Number of client models in $c$ cluster)
    $\omega^c_{t+1} \leftarrow \frac{1}{|m|}\sum_m \omega^{c,m}_{t+1}$
  **end for**
  $\omega_{t+1} \leftarrow \frac{1}{|C_{t+1}|}\sum_c \omega^c_{t+1}$
**end for**

#### 4.4.1. Model training

Unlike the FedAvg algorithm mentioned in related works, our scheme begins with the server performing the initial model training. The pre-trained model is then sent to participating clients who use it to train their local models. The loss function used in the local model training is detailed in the preliminaries section. Finally, clients send their local model parameters back to the server, completing the model training round.

#### 4.4.2. Model aggregation

In the ECG monitoring system, the distribution of data among clients may not be independent and homogeneous, which can lead to to the complex convergence of the model aggregation during federated learning. Therefore, it's crucial to address the non-i.i.d. problem appropriately in widely used FedAvg and FedProx model aggregation algorithms. To solve this problem, we propose a model weights clustering algorithm performed on the server to update model aggregation efficiently.

The precise steps of the model weights clustering algorithm are shown in Algorithm 2. We use the K-means algorithm as the ClusteringAlgorithm. To select the clustering point $k$, we determine the best value by running each local model parameter through several K-means clustering runs.

## 5. Experiments

In this section, we conduct extensive experiments on a real-world ECG dataset to demonstrate the effectiveness of our FedECG framework. Firstly, we assess the efficacy of our suggested framework and compare it with alternative approaches using default parameters. Secondly, we investigate the impact of various parameters, specifically to evaluate the efficiency of the presented model clustering aggregation algorithm in the non-i.i.d. scenario. Lastly, we examine the operational efficiency and communication costs of the proposed framework.

### 5.1. Dataset description

To create a realistic experimental setting, we use the raw ECG dataset MIT-BIH from the PhysioNet website, currently managed by the MIT Laboratory for Computational Physiology (Goldberger et al., 2000). The dataset contains ECG recordings from more than 40 patients spanning five years, comprising around ten thousand ECG beats with five different arrhythmia categories in compliance with the AAMI EC57 standard. Before the experiment, we preprocess the dataset and use its labeled data as real data. Ultimately, we obtain a total of 109,446 samples. We allocate 10,944 samples as the test set and distribute the remainder proportionally to the labeled data as the server and client datasets. Importantly, we remove the labels from the client dataset to simulate real-world application scenarios.

### 5.2. Experimental setup

We specify the following default values for some common parameters in our simulation experiments:

- **Proportion of labeled data** $\gamma$: $\gamma$ denotes the percentage of labeled data available to the server. We set the initial value of $\gamma$ to 50%.
- **Number of total clients** $N$: $N$ represents the total number of clients, each with their own unlabeled data for model training. The default value for $N$ is 100.
- **Fraction of active clients** $K$: Only a subset of clients participate in each communication round of federated learning. We set the default value of $K$ to 20.
- **Number of communication rounds** $T$: $T$ indicates the total number of rounds of model training. We set the initial value of $T$ to 100.
- **Level of non-i.i.d.** $Q$: Each client only possesses data for certain types of ECG in real scenarios. Generally, healthy individuals have data for only one type, while those with cardiovascular diseases may have data for two or more types. To simulate such non-i.i.d. scenarios, we set the value of $Q$ to 0.5.

Each client utilizes the SGD optimizer for model training with a local batch size of ten in each round. After five training epochs by the clients, the server performs model aggregation using the model weights clustering algorithm. We conduct all experiments using Tensorflow and Pytorch and simulate the experiments on a PC equipped with an Intel(R) i5 12600 K CPU and an Nvidia GeForce RTX1060 6G GPU.

**Table 3**
Comparison of the efficiency of ECG abnormalities prediction methods.

| Method | Accuracy (%) | FL |
|---|---|---|
| CNN (Acharya et al., 2017) | 93.5 | ✗ |
| SVM (Martis et al., 2013) | 93.8 | ✗ |
| Random Forest (Li and Zhou, 2016) | 94.6 | ✗ |
| LSTM (Wang et al., 2022) | 92 | ✓ |
| ResNet-18 (Jing et al., 2021) | 96.5 | ✗ |
| ResNet-9 (Ours) | 95.9($\gamma$=100%) 94.8($\gamma$=50%) | ✓ |

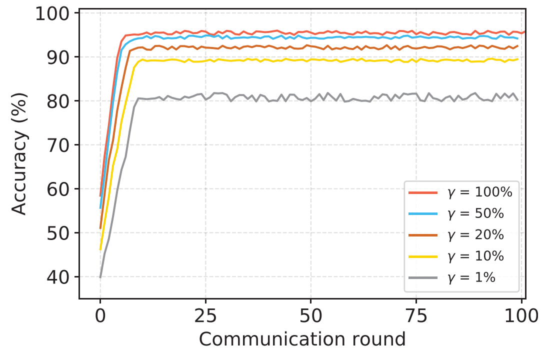


**Fig. 6.** Model performance at different proportions of labeled data.

### 5.3. Model performance

To evaluate the impact of the proportion of labeled data on the framework's performance, we conduct experiments with different percentages, including 1%, 10%, 20%, 50%, and 100%. The results are presented in Table 3, where we can observe that our proposed framework achieves an accuracy of 94.8% with only 50% labeled data, which is close to the fully supervised learning situation with 95.9% accuracy. Moreover, as shown in Fig. 6, the framework still performs well with an accuracy rate of 85% when only 1% of the data is labeled. Our model can converge quickly in fewer than ten communication rounds, reaching stable accuracy and minimal loss after 100 communication rounds. It is worth noting that using less labeled data also means better protection of user privacy, as less data is revealed.

We also evaluate the performance of our model in non-i.i.d. scenarios with real ECG applications by setting different levels of non-i.i.d. Our model performs slightly worse under non-i.i.d. but still achieves high accuracy compared to i.i.d. The accuracy varies within a 5% interval for various data distributions, indicating that our framework is also friendly to non-i.i.d. scenarios.

### 5.4. Performance against other models

We evaluate the performance of our proposed FedECG framework in comparison with the prior studies. FedECG achieves comparable accuracy to other methods while addressing privacy-preserving and label scarcity concerns that are not tackled by other methods. We compare our model with other methods, including CNN (Acharya et al., 2017), SVM (Martis et al., 2013), Random Forest (RF) (Li and Zhou, 2016), LSTM (Wang et al., 2022), and ResNet-18 (Jing et al., 2021). All of these methods are trained on the MIT-BIH ECG dataset, which we also utilize.

The comparison values are given in Table 3, demonstrating that our model achieved a high accuracy among the evaluated approaches. Our FedECG model outperforms the conventional CNN model, even when trained on only 50% of the labeled data. However, when utilizing a smaller amount of labeled data, our model's accuracy falls below that of the LSTM, dropping to less than 90%. Although semi-supervised learning can provide pseudo labels, it may also generate false ones, leading to reduce model accuracy. Moreover, the model's performance in an FL setting can be further impacted by the presence of non-i.i.d. data. Although ResNet-18 exhibits higher accuracy than our model, our model architecture has fewer layers and is better suited for IoT smart devices with limited computational power. The average communication round time for one client in our simulation experiments is approximately 4.3 s, significantly decreasing communication costs. Most importantly, our framework does not require a centralized server to learn user data, thereby protecting user privacy. Therefore, our approach optimizes the efficiency and accuracy of model training in scenarios that prioritize user privacy and cope with label scarcity.

### 5.5. Performance under different factors

We assess the impact of several factors on the framework's performance, including the model aggregation algorithm, the number of selected clients, the number of communication rounds, and the level of non-i.i.d. data. By analyzing these variables, we aim to gain a deeper understanding of how each aspect contributes to the overall effectiveness of the system, ultimately providing insights for optimizing the framework's performance.

#### 5.5.1. Impact of the model algorithm

To evaluate the efficacy of the model weights clustering algorithm in tackling non-i.i.d. data, we conduct simulations using various non-i.i.d. levels. We set $\gamma$ to 0.5 while maintaining other parameters at their initial values. The results are illustrated in Fig. 7. The accuracy achieved with the model weights aggregation algorithm is 2.7% higher than that obtained without it. Additionally, the clustering algorithm reduces the model convergence time by approximately ten communication rounds. The server initially groups the model gradients for each client to minimize the impact of non-i.i.d. data and accelerate model aggregation. Consequently, our proposed model weights clustering algorithm effectively addressesthe challenge posed by non-i.i.d. data.

#### 5.5.2. Impact of the client number

We examine the impact of varying client numbers on the efficiency of our framework. For the simulation, we consider the

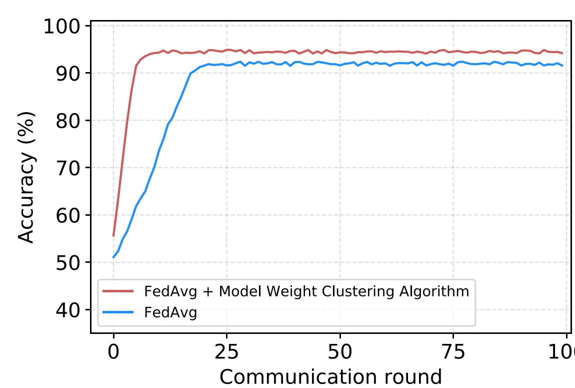


**Fig. 7.** Accuracy comparison with or without the model weights clustering algorithm.

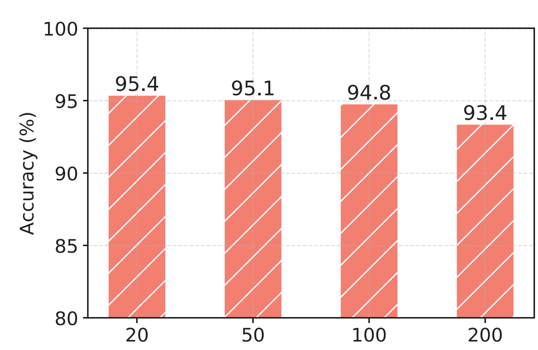


**Fig. 8.** Accuracy comparison with different numbers of clients.

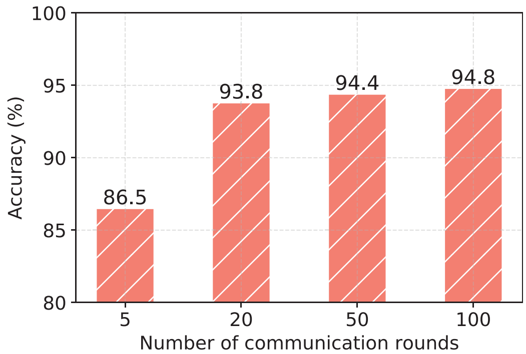


**Fig. 9.** Accuracy comparison with different numbers of communication rounds.

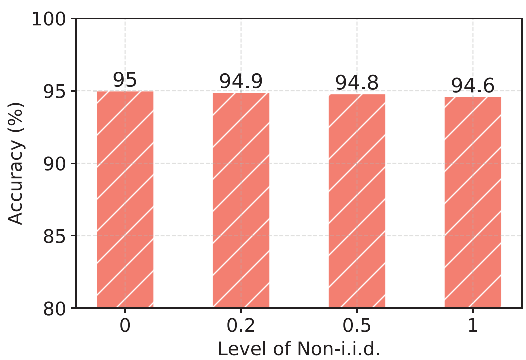


**Fig. 10.** Accuracy comparison with different levels of non-i.i.d.

number of clients $N = 20, 50, 100, 200$ and the number of active clients $K = 4, 10, 20, 40$. As depicted in Fig. 8, there is a tendency for the framework's performance to decrease with an increasing number of clients. This occurs because a higher number of clients leads to greater data diversity, which can result in increased non-i.i.d. data and less effective local models for aggregated clients. While an increased number of clients can diminish the framework's performance, this is not our primary concern. This work still chooses the total number of clients $N = 100$ to simulate a realistic ECG monitoring scenario.

### 5.5.3. Impact of the communication round

We evaluate the impact of different numbers of communication rounds $T$ on the framework's performance. Specifically, we consider $T = 5, 20, 50, 100$ and measure the framework's accuracy. As shown in Fig. 9, the framework's performance generally improves as communication rounds increase. Greater communication between the server and clients leads to better aggregation of local models and more accurate global models. However, we observe that the performance improvement diminishes after approximately 25 rounds, indicating that the framework converges quickly and further communication offers diminishing returns. Therefore, in practical settings, setting the number of communication rounds to around 25 can strike a balance between reducing communication overhead and maintaining high model accuracy.

### 5.5.4. Impact of the Non-i.i.d. level

We conduct an experiment to investigate the impact of different levels of non-i.i.d. data on the performance of our framework model. We set the non-i.i.d. level $\gamma$ from 0 to 1, where a value of 1 means that each client has only one type of dataset. The results are shown in Fig. 10. We note that increasing levels of non-i.i.d. data have no significant effect on the framework's performance. This can be attributed to our proposed model clustering aggregation algorithm, which effectively groups non-independent and homogeneously distributed data for robust aggregation. This result demonstrates the resilience of our framework in handling various levels of non-i.i.d. data.

## 6. Conclusion

In this work, we present a federated semi-supervised learning framework for predicting ECG abnormalities while protecting user privacy and addressing label scarcity. To process the noisy ECG data, we suggest using wavelet transformation and converting the one-dimensional ECG signals into two-dimensional ECG images using Gramian Angular Summation/Difference Fields for feature extraction. Considering the computational constraints of ECG devices, we design a novel ResNet-9 model with fewer layers and faster convergence. To increase the amount of unlabeled data available to clients, we employ pseudo-labeling and data augmentation techniques for semi-supervised learning. Benefiting from federated learning, the continuously updated iterative global model enhances the accuracy of the pseudo labels, rendering the unlabeled data useful. To address the challenge of non-i.i.d. data gathered from clients, we propose a model weight clustering algorithm based on FedAvg for robust aggregation.

We conduct several simulations on the MIT-BIH dataset to assess the performance of the proposed framework. Initially, we evaluate the framework's performance based on the percentage of differently labeled data. Our FedECG achieves 95.9% accuracy with supervised learning using 100% labeled data and 94.8% accuracy with only half of the labeled data. Although the accuracy is similar to previous methods, our proposed framework protects data privacy and solves the issue of missing annotated data. We also examine the performance of our framework under various parameter setups, such as the level of non-i.i.d. distribution, the number of clients, and the number of communication rounds. We find that our framework is efficient in these situations and generates robust aggregation with our model clustering algorithm. This work presents a promising approach for the prediction of cardiovascular diseases on a large scale and has the potential to be extended to other health diagnostics, such as blood glucose and blood pressure monitoring.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.


## Acknowledgement

This research is supported by FDCT under its General R&D Subsidy Program Fund (Grant No. 0038/2022/A), Macau.


## References


Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M., Gertych, A., San Tan, R., 2017. A deep convolutional neural network model to classify heartbeats. Comput. Biol. Med. 89, 389–396.

Afkhami, R.G., Azarnia, G., Tinati, M.A., 2016. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. Pattern Recognit. Lett. 70, 45–51.

Ali, F., El-Sappagh, S., Islam, S.R., Ali, A., Attique, M., Imran, M., Kwak, K.S., 2021. An intelligent healthcare monitoring framework using wearable sensors and social networking data. Future Generat. Comput. Syst. 114, 23–43.

Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M., Kwak, K.S., 2020. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Informat. Fus. 63, 208–222.

Almotairi, K.H., Hussein, A.M., Abualigah, L., Abujayyab, S.K., Mahmoud, E.H., Ghanem, B.O., Gandomi, A.H., 2023. Impact of artificial intelligence on covid-19 pandemic: A survey of image processing, tracking of disease, prediction of outcomes, and computational medicine. Big Data Cognitive Comput. 7, 11.
Baig, M.M., Gholamhosseini, H., Connolly, M.J., 2013. A comprehensive survey of wearable and wireless ecg monitoring systems for older adults. Medical Biol. Eng. Comput. 51, 485–495.
Barstuğan, M., Ceylan, R., 2020. The effect of dictionary learning on weight update of adaboost and ecg classification. J. King Saud Univ.- Comput. Informat. Sci. 32, 1149–1157.
Bin, Y., Cao, X., Chen, X., Ge, Y., Tai, Y., Wang, C., Li, J., Huang, F., Gao, C., Sang, N., 2020. Adversarial semantic data augmentation for human pose estimation. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIX 16. Springer, pp. 606–622.
Chatrati, S.P., Hossain, G., Goyal, A., Bhan, A., Bhattacharya, S., Gaurav, D., Tiwari, S. M., 2022. Smart home health monitoring system for predicting type 2 diabetes and hypertension. J. King Saud Univ.- Comput. Informat. Sci. 34, 862–870.
Choi, T., Would, O., Salazar-Gomez, A., Cielniak, G., 2022. Self-supervised representation learning for reliable robotic monitoring of fruit anomalies. In: 2022 International Conference on Robotics and Automation (ICRA). IEEE, pp. 2266–2272.
Ebrahimi, Z., Loni, M., Daneshtalab, M., Gharehbaghi, A., 2020. A review on deep learning methods for ecg arrhythmia classification. Expert Syst. Appl.: X 7, 100033.
El-Sappagh, S., Ali, F., Abuhmed, T., Singh, J., Alonso, J.M., 2022. Automatic detection of alzheimer's disease progression: An efficient information fusion approach with heterogeneous ensemble classifiers. Neurocomputing 512, 203–224.
van Engelen, J.E., Hoos, H.H., 2020. A survey on semi-supervised learning. Mach. Learn. 109, 373–440.
Golany, T., Freedman, D., Radinsky, K., 2021. Ecg ode-gan: Learning ordinary differential equations of ecg dynamics via generative adversarial learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 134–141.
Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E., 2000. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation 101, e215–e220.
Guddhur Jayadev, P., Bellary, S., 2022. A hybrid approach for classification and identification of iris damaged levels of alcohol drinkers. J. King Saud Univ.- Comput. Informat. Sci. 34, 5273–5285.
Guo, R., Peng, Y., Kong, W., Li, F., 2022. A semi-supervised label distribution learning model with label correlations and data manifold exploration. J. King Saud Univ.- Comput. Informat. Sci. 34, 10094–10108.
Hartigan, J.A., Wong, M.A., 1979. Algorithm as 136: A k-means clustering algorithm. J. Roy. Stat. Soc.. Series C (Appl. Stat.) 28, 100–108.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.
Huang, P., Guo, L., Li, M., Fang, Y., 2019. Practical privacy-preserving ecg-based authentication for iot-based healthcare. IEEE Internet Things J. 6, 9200–9210.
Hussein, F., Mughaid, A., AlZu'bi, S., El-Salhi, S.M., Abuhaija, B., Abualigah, L., Gandomi, A.H., 2022. Hybrid clahe-cnn deep neural networks for classifying lung diseases from x-ray acquisitions. Electronics 11, 3075.
Jing, E., Zhang, H., Li, Z., Liu, Y., Ji, Z., Ganchev, I., 2021. Ecg heartbeat classification based on an improved resnet-18 model. Comput. Mathe. Methods Med. 2021.
Kachuee, M., Fazeli, S., Sarrafzadeh, M., 2018. Ecg heartbeat classification: A deep transferable representation. In: 2018 IEEE international Conference on Healthcare Informatics (ICHI). IEEE, pp. 443–444.
Kachuee, M., Kiani, M.M., Mohammadzade, H., Shabany, M., 2017. Cuffless blood pressure estimation algorithms for continuous health-care monitoring. IEEE Trans. Biomed. Eng. 64, 859–869.
Kaptoge, S., Pennells, L., De Bacquer, D., Cooney, M.T., Kavousi, M., Stevens, G., Riley, L.M., Savin, S., Khan, T., Altay, S., et al., 2019. World health organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Global Health 7, e1332–e1345.
Kocabas, O., Soyata, T., 2020. Towards privacy-preserving medical cloud computing using homomorphic encryption. In: Virtual and Mobile Healthcare: Breakthroughs in Research and Practice. IGI Global, pp. 93–125.
Lee, D.H. et al., 2013. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In: Workshop on Challenges in Representation Learning. ICML, p. 896.
Li, J., Pang, S.p., Xu, F., Ji, P., Zhou, S., Shu, M., 2022. Two-dimensional ecg-based cardiac arrhythmia classification using dse-resnet. Sci. Rep. 12, 14485.
Li, T., Zhou, M., 2016. Ecg classification using wavelet packet entropy and random forests. Entropy 18, 285.
Martis, R.J., Acharya, U.R., Lim, C.M., Mandana, K., Ray, A.K., Chakraborty, C., 2013. Application of higher order cumulant features for cardiac health diagnosis using ecg signals. Int. J. Neural Syst. 23, 1350014.
McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A., 2017. Communication-efficient learning of deep networks from decentralized data. In: Artificial Intelligence and Statistics. PMLR, pp. 1273–1282.
Mehari, T., Strodthoff, N., 2022. Self-supervised representation learning from 12-lead ecg data. Comput. Biol. Med. 141, 105114.
Moody, G.B., Mark, R.G., 2001. The impact of the mit-bih arrhythmia database. IEEE Eng. Med. Biol. Mag. 20, 45–50.
Moody, G.B., Mark, R.G., Goldberger, A.L., 2001. Physionet: a web-based resource for the study of physiologic signals. IEEE Eng. Med. Biol. Mag. 20, 70–75.
Pandey, B., Kumar Pandey, D., Pratap Mishra, B., Rhmann, W., 2022. A comprehensive survey of deep learning in the field of medical imaging and medical natural language processing: Challenges and research directions. J. King Saud Univ.- Comput. Informat. Sci. 34, 5083–5099.
Philip, N.Y., Rodrigues, J.J., Wang, H., Fong, S.J., Chen, J., 2021. Internet of things for in-home health monitoring systems: current advances, challenges and future directions. IEEE J. Sel. Areas Commun. 39, 300–310.
Raza, A., Tran, K.P., Koehl, L., Li, S., 2022. Designing ECG monitoring healthcare system with federated transfer learning and explainable AI. Knowl. Based Syst. 236, 107763.
Regulation, P., 2016. Regulation (eu) 2016/679 of the european parliament and of the council. Regulation (eu) 679, 2016.
Rodrigues, R., Couto, P., 2021. Semi-supervised learning for ecg classification. In: 2021 Computing in Cardiology (CinC). IEEE, pp. 1–4.
Roth, G.A., Mensah, G.A., Johnson, C.O., Addolorato, G., Ammirati, E., Baddour, L.M., Barengo, N.C., Beaton, A.Z., Benjamin, E.J., Benziger, C.P., et al., 2020. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the gbd 2019 study. J. Am. Coll. Cardiol. 76, 2982–3021.
Saadatnejad, S., Oveisi, M., Hashemi, M., 2019. Lstm-based ecg classification for continuous monitoring on personal wearable devices. IEEE J. Biomed. Health Informat. 24, 515–523.
Sabor, N., Gendy, G., Mohammed, H., Wang, G., Lian, Y., 2022. Robust arrhythmia classification based on qrs detection and a compact 1d-cnn for wearable ecg devices. IEEE J. Biomed. Health Informat.
Şahinbaş, K., Catak, F.O., 2021. Secure multi-party computation based privacy preserving data analysis in healthcare iot systems. arXiv preprint arXiv:2109.14334.
Sandfort, V., Yan, K., Pickhardt, P.J., Summers, R.M., 2019. Data augmentation using generative adversarial networks (cyclegan) to improve generalizability in ct segmentation tasks. Sci. Rep. 9, 16884.
Sarkar, P., Etemad, A., 2020. Self-supervised learning for ecg-based emotion recognition. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 3217–3221.
Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6, 1–48.
Singh, M., Pujar, G.V., Kumar, S.A., Bhagyalalitha, M., Akshatha, H.S., Abuhaija, B., Alsoud, A.R., Abualigah, L., Beeraka, N.M., Gandomi, A.H., 2022. Evolution of machine learning in tuberculosis diagnosis: a review of deep learning-based medical applications. Electronics 11, 2634.
Stergiou, G.S., Alpert, B., Mieke, S., Asmar, R., Atkins, N., Eckert, S., Frick, G., Friedman, B., Graßl, T., Ichikawa, T., et al., 2018. A universal standard for the validation of blood pressure measuring devices: Association for the advancement of medical instrumentation/european society of hypertension/ international organization for standardization (aami/esh/iso) collaboration statement. Hypertension 71, 368–374.
Tuncer, T., Dogan, S., Pławiak, P., Acharya, U.R., 2019. Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ecg signals. Knowl.-Based Syst. 186, 104923.
Wagan, S.A., Koo, J., Siddiqui, I.F., Attique, M., Shin, D.R., Qureshi, N.M.F., 2022. Internet of medical things and trending converged technologies: A comprehensive review on real-time applications. J. King Saud Univ.- Comput. Informat. Sci. 34, 9228–9251.
Wagner, T., Guha, S., Kasiviswanathan, S., Mishra, N., 2018. Semi-supervised learning on data streams via temporal label propagation. In: International Conference on Machine Learning, PMLR. pp. 5095–5104.
Wang, X., Hu, J., Lin, H., Liu, W., Moon, H., Piran, M.J., 2022. Federated learning-empowered disease diagnosis mechanism in the internet of medical things: From the privacy-preservation perspective. IEEE Trans. Ind. Informat.
Wang, Z., Oates, T., 2015. Imaging time-series to improve classification and imputation. In: Proceedings of the 24th International Conference on Artificial Intelligence. AAAI Press, pp. 3939–3945.
Wu, M., Lu, Y., Yang, W., Wong, S.Y., 2020. A study on arrhythmia via ECG signal classification using the convolutional neural network. Frontiers Comput. Neurosci. 14, 564015.
Wyatt, K.D., Poole, L.R., Mullan, A.F., Kopecky, S.L., Heaton, H.A., 2020. Clinical evaluation and diagnostic yield following evaluation of abnormal pulse detected using apple watch. J. Am. Med. Inform. Assoc. 27, 1359–1363.
Yang, C.Y., Wang, W.F., 2022. An efficient data hiding for ecg signals based on the integer wavelet transform and block standard deviation. J. King Saud Univ.- Comput. Informat. Sci. 34, 5868–5882.
Yang, Q., Liu, Y., Chen, T., Tong, Y., 2019. Federated machine learning: Concept and applications. ACM Trans. Intell. Syst. Technol. (TIST) 10, 1–19.
Yang, Z., Zhou, Q., Lei, L., Zheng, K., Xiang, W., 2016. An iot-cloud based wearable ecg monitoring system for smart healthcare. J. Medical Syst. 40, 1–11.
Ying, Z., Cao, S., Liu, X., Ma, Z., Ma, J., Deng, R.H., 2022. Privacysignal: Privacy-preserving traffic signal control for intelligent transportation system. IEEE Trans. Intell. Transp. Syst. 23, 16290–16303.
Ying, Z., Cao, S., Xu, S., Liu, X., Lyu, L., Chen, C., Wang, L., 2021. Privacy-preserving optimal insulin dosing decision. In: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 2640–2644.
Yuan, J., Liu, Y., Shen, C., Wang, Z., Li, H., 2021. A simple baseline for semi-supervised semantic segmentation with strong data augmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 8229–8238.
Yusuf, S., Joseph, P., Rangarajan, S., Islam, S., Mente, A., Hystad, P., Brauer, M., Kutty, V.R., Gupta, R., Wielgosz, A., et al., 2020. Modifiable risk factors, cardiovascular

disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (pure): a prospective cohort study. The Lancet 395, 795–808.
Zhai, X., Zhou, Z., Tin, C., 2020. Semi-supervised learning for ECG classification without patient-specific labeled data. Expert Syst. Appl. 158, 113411.
Zhang, D., 2018. Big data security and privacy protection. In: 8th International Conference on Management and Computer Science (ICMCS 2018). Atlantis Press, pp. 275–278.
Zhang, M., Wang, Y., Luo, T., 2020. Federated learning for arrhythmia detection of non-iid ecg. In: 2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, pp. 1176–1180.
Zhang, Y., Li, J., Wei, S., Zhou, F., Li, D., 2021. Heartbeats classification using hybrid time-frequency analysis and transfer learning based on resnet. IEEE J. Biomed. Health Informatics 25, 4175–4184.
Zhang, Y., Wei, X., Cao, J., Ning, J., Ying, Z., Zheng, D., 2022. Blockchain-enabled decentralized attribute-based access control with policy hiding for smart healthcare. J. King Saud Univ.-Comput. Informat. Sci. 34, 8350–8361.
Zhang, Y., Ying, Z., Chen, C.P., 2022. Achieving privacy-preserving multitask allocation for mobile crowdsensing. IEEE Internet Things J. 9, 16795–16806.
Zheng, M., You, S., Wang, F., Qian, C., Zhang, C., Wang, X., Xu, C., 2021. Ressl: Relational self-supervised learning with weak augmentation. Adv. Neural Informat. Process. Syst. 34, 2543–2555.